Docket No. TI-37043 PATENT APPLICATION

MINIMIZING TRANSISTOR VARIATIONS DUE TO SHALLOW TRENCH ISOLATION STRESS

Inventor: Jong Shik Yoon
3604 Snidow Drive
Plano, Tx 75025

Andrew Tae Kim
3800 Pebble Creek Court #1213
Plano, Tx 75023

Assignee: Texas Instruments, Incorporated
P.O. Box 655474, MS 3999
Dallas, TX 75256

CERTIFICATE OF EXPRESS MAIL

I hereby certify that this correspondence, including the attachments listed, is being deposited with the United States Postal Service, Express Mail - Post Office to Addressee, Receipt No. EV 372643555 US, in an envelope addressed to Commissioner of Patents and Trademarks, Washington, D.C. 20231, on the date shown below.

4-15-2009 Elizabeth Schumacher
Date of Mailing Typed or printed name of person mailing

Elizabeth Schumacher
Signature of person mailing

Hitt Gaines, P.C.
P.O. Box 832570
Richardson, Texas 75083
(972) 480-8800

MINIMIZING TRANSISTOR VARIATIONS DUE TO SHALLOW TRENCH ISOLATION STRESS

TECHNICAL FIELD OF THE INVENTION

[0001] The present invention is directed, in general, to the manufacture of semiconductor devices.

BACKGROUND OF THE INVENTION

[0002] The continuing push to produce faster semiconductor devices with lower power consumption has resulted in transistor miniaturization and higher integrated circuit packing densities with each new technology node. Moreover, transistors with smaller gate dimensions and channel width and higher packing density are conducive to faster operating devices. Along with shrinking transistor geometries, however, comes a number of challenges to optimize both transistor design and integrated circuit (IC) layout design.

[0003] IC layout design involves determining the optimal placement and packing of active and passive areas over a silicon substrate, typically with the assistance of a ruled-based computer-aided designs (CAD) program. Active areas are defined as regions of the silicon substrate on which operative components are built, such as transistors, capacitors and resistors. Some active areas,

also known as diffusion regions, are n- or p-doped to form transistors. To prevent conduction or crosstalk between active areas, the individual active areas are bounded by passive areas. Passive areas, also referred to as isolation structures, are usually formed by filling trenches etched in the substrate with an insulator, such as silicon dioxide. In addition, other non-operative areas of the silicon substrate, known as active dummies, which may not be doped, are distributed through the IC layout to help ensure that a planar surface is formed during chemical mechanical polishing.

[0004] Numerous design rules are used to decide the dimension and relative locations of active areas, passive areas and active dummies. For instance, transistors are designed by overlapping an active area and a polysilicon layer. The polysilicon layer can be used to form the gate electrode for one or more transistors, as well as to interconnect layers. Conventional design rules consider the minimum amount that the gate must overlap the active area ("gate overlap"), and the minimum amount that the active area must overlap the gate ("active overlap"). The rule for minimum active overlap ensures that there is an adequate area to form contacts to source/drain structures in the active area, and to maintain electrical isolation between the contacts and the gate. Still other design rules, based on the transistor's channel length and width dimensions, are used to predict the transistor's performance

characteristics, such as the on current (I_{on}), threshold voltage (V_{th}) and off drive current (I_{off}). For example, the transistor's expected performance may be modified to better interface with other active areas by adjusting the channel length and width, by adjusting the gate's dimensions.

[0005] Conventional design rules also ensure that active dummies are not too close to active areas. As an example, current technology nodes may call for a minimum separation distance between active areas and passive areas of between 1000 and 2000 nanometers. In addition, there are also design rules to restrict the placement of active dummies near polysilicon or other gate structures, to reduce parasitic capacitance. Parasitic capacitance can be further reduced by restricting the shape of active dummies to be pillars instead of one large continuous block. Typically, active dummies have variable sizes, shapes and locations relative to active areas for different areas of the substrate.

[0006] Unfortunately, conventional design rules can produce transistors whose predicted and actual performance characteristics differ markedly. Moreover, the trend is for the disparity between predicted and actual performance to increase as the transistor's dimensions are scaled down and circuit density is increased. A deviation between predicted and actual performance can cause the IC to malfunction. To correct these deviations, it may be necessary to redesign the IC layout. These complexities, in turn, can

increase manufacturing costs and delay the production of ICs.

[0007] Accordingly, what is needed in the art is an improved method of manufacturing transistors and integrated circuits with predictable performance characteristics and high transistor packing densities, while not suffering the deficiencies of previous approaches.

SUMMARY OF THE INVENTION

[0008] To address the above-discussed deficiencies of the prior art, one embodiment of the present invention is a method of manufacturing a metal oxide semiconductor (MOS) transistor. The method includes forming an active area in a silicon substrate. The active area is bounded by an isolation structure. The method also includes placing at least one stress adjustor adjacent the active area. The stress adjustor is positioned to modify a mobility of a majority carrier within a channel region of the MOS transistor.

[0009] Another embodiment of the present invention is a MOS transistor. The MOS transistor comprises an active area in a substrate and an isolation structure. The isolation structure is in the substrate and surrounding the active area. The MOS transistor also includes at least one stress adjustor adjacent the active area. The stress adjustor is positioned to modify a mobility of a majority carrier within a channel region of the MOS transistor.

[0010] Still another embodiment of the present invention is a method of constructing an integrated circuit (IC). The method includes generating a mask layout for the IC and using the mask layout to fabricate the IC. Generating the mask layout includes calculating an active overlap distance between a planned perimeter of a gate of the IC and a planned perimeter of an active area of

the IC. A compressive stress along a direction of an intended current flow through a planned channel of said IC is determined based on the active overlap distance. A stress adjustor area adjacent the active area is introduced to modify a mobility of a majority carrier through the planned channel, if the compressive stress is greater than a critical stress parameter.

[0011] The foregoing has outlined preferred and alternative features of the present invention so that those of ordinary skill in the art may better understand the detailed description of the invention that follows. Additional features of the invention will be described hereinafter that form the subject of the claims of the invention. Those skilled in the art should appreciate that they can readily use the disclosed conception and specific embodiment as a basis for designing or modifying other structures for carrying out the same purposes of the present invention. Those skilled in the art should also realize that such equivalent constructions do not depart from the scope of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] The invention is best understood from the following detailed description when read with the accompanying FIGUREs. It is emphasized that in accordance with the standard practice in the semiconductor industry, various features may not be drawn to scale. In fact, the dimensions of the various features may be arbitrarily increased or reduced for clarity of discussion. Reference is now made to the following descriptions taken in conjunction with the accompanying drawings, in which:

[0013] FIGUREs 1A to 1N illustrate a partial sectional and plan views of selected steps in a method for manufacturing a MOS device according to the principles of the present invention;

[0014] FIGUREs 2A and 2B illustrate sectional and plan views, respectively, of an exemplary MOS device of the present invention; and

[0015] FIGURE 3 illustrates by flow diagram, a method for constructing an integrated circuit layout according to the principles of the present invention.

DETAILED DESCRIPTION

[0016] The present invention recognizes for the first time that as the transistor device dimensions are scaled down, compressive stress from the isolation structure surrounding a transistor can substantially affect the mobility of majority carriers through the transistor's channel region. For NMOS transistors, compressive stress in the direction of current flow through the channel can decrease the mobility of electrons in the channel. Conversely, for PMOS transistors, compressive stress in the direction of current flow can enhance the mobility of holes through the channel.

[0017] This recognition, in turn, lead to the realization that active overlap is an important design parameter affecting the performance characteristics of transistors. For transistors whose active overlap reaches a certain critical small value, the amount of stress imparted to the transistor's channel can be advantageously modified. In particular, stress adjusting structure can be introduced adjacent the active area, to either increase or decrease the compressive stress to adjust channel mobility as desired.

[0018] One embodiment of the present invention is a method for manufacturing a MOS device. FIGUREs 1A to 1N illustrate partial sectional and plan views of selected steps in an exemplary method of manufacturing a Metal Oxide Semiconductor (MOS) transistor 100.

Turning first to FIGURE 1A, illustrated is a plan view the partially completed MOS transistor 100 while FIGURE 1B present a cross-sectional view of the same structure, corresponding to view A-A in FIGURE 1A. Analogous views are presented in FIGUREs 1C-1N below.

[0019] FIGUREs 1A-1D illustrate the formation of an active area. As shown in FIGUREs 1A and 1B, defining an active area 105 includes forming a trench 110 in a substrate 115. The substrate 115 preferably is comprised of silicon. Those skilled in the art would be familiar with the conventional lithographic processes used to form trenches in silicon including covering portions of silicon substrate 115 with a photoresist and wet or plasma etching the uncovered portions to form. FIGUREs 1C and 1D show the partially completed MOS transistor 100 after using conventional procedures to fill the trench 110 with an insulating material, such as silicon dioxide, to form an isolation structure 120. The active area 105 is thereby bounded by the isolation structure 120, as represented by the dashed lines 122.

[0020] FIGUREs 1A-1D also illustrate a preferred method to place a stress adjustor adjacent the active area. FIGUREs 1A and 1B illustrate that forming the trench 110 includes removing portions of the silicon substrate 115 to form at least two trenches 125 to define a stress adjustor 130. As shown in FIGUREs 1C and 1D, the two trenches 125 are filled with the insulating material of the

isolation structure 120. The portion of silicon remaining between the trenches 125 is the stress adjustor 130.

[0021] FIGUREs 1E-1L illustrate an alternative method to place a stress adjustor adjacent the active area. As shown in FIGUREs 1E and 1F, a trench 110 is formed in a silicon substrate 115. Then, as shown in FIGUREs 1G and 1H, the trench 110 is filled with an insulating material to form an isolation structure 120, thereby defining the boundary 122 around the active area 105. Referring now to FIGUREs 1I and 1J, a portion of the isolation structure 120 is removed using conventional procedures to form a trench 135 adjacent the active area 105, and preferably within the isolation structure 120. Next, as illustrated in FIGUREs 1K and 1L, the trench 135 is filled with a material to form the stress adjustor 130.

[0022] With continuing reference to FIGURE 1I, in some preferred embodiments, the trench 135 is filled with an insulating material comprising silicon dioxide via a high density plasma (HDP) oxide deposition. HDP oxide deposition is preferred when the trench 135 has a narrow width 140, because this method is conducive to filling narrow structures (*e.g.*, width 140 of less than about 150 nanometers). In other instances, however, other methods can be used to fill the trench 135, such as spin-on-glass type oxides.

[0023] Turning now to FIGUREs 1M and 1N, shown is the partially constructed MOS transistor 100 of FIGUREs 1C and 1D, after forming

a gate 150 over the active area 105. Conventional oxidation, deposition and lithographic procedures can be used to deposit and pattern oxide and polysilicon layers to form the insulator and electrode of the gate 150. Of course, other well-known materials may also be included in the formation of the gate 150. The dimensions of the gate 150 define the dimensions of a channel region 155 in the active area 105 that is under the gate 150.

[0024] With continuing reference to FIGUREs 1A-1N, the material used to form the stress adjustor 130 is selected either to increase or decrease a compressive stress imparted from the isolation structure 120 to the channel region 155. For example, when the MOS transistor 100 is an NMOS transistor, it is desirable to decrease the compressive stress in an intended direction of current flow 160.

[0025] Although not depicted in FIGUREs 1A-1N, one skilled in the art would understand that the manufacture of the MOS transistor 100 could also include conventional steps to form other transistor components, such as doped wells, source/drain and source/drain extensions and gate sidewall structures.

[0026] Another embodiment of the present invention is a MOS transistor. FIGUREs 2A and 2B, present plan and cross-sectional views, respectively, of an exemplary MOS transistor 200 of the present invention. Any of the above-described embodiments of the methods of manufacturing the MOS device 100, such as depicted in

FIGUREs 1A-1N, can be used to construct the MOS transistor 200.

[0027] The MOS transistor 200 comprises an active area 205 in a substrate 210 and an isolation structure 215. The isolation structure 215 is in the substrate 210 and surrounds the active area 205. For each technology node, minimum design rules specify the minimum dimensions of the isolation structure 215 needed to isolate the active area 205 from adjacent active areas (not shown). In some embodiments for instance, the isolation structure 215 is a shallow trench isolation (STI) structure, having a width 217 of at least about 50 to about 200 nanometers and a depth 219 of at least about 250 to about 400 nanometers. Of course, other shapes, widths 215 and depths 219 for the isolation structure 215 are also within the scope of the present invention.

[0028] Preferred embodiments of the MOS transistor 200 also include a gate 220 located over the active area 205. A channel region 225 is located in the active area 205 under the gate 220. Similar to the isolation structure 215, design rules, well known to those skilled in the art, govern the dimensions of the gate 220 according to the desired performance characteristics of the transistor 200 for the technology node of interest. Typically, a gate length 227 is fixed for a particular technology node (e.g., 0.25 microns) and a gate width 228 is adjusted by the circuit designer to obtain the desired current.

[0029] The MOS transistor 200 also includes at least one stress

adjustor 230 adjacent the active area 205. When the MOS transistor 200 includes a plurality of stress adjustors 230, the individual stress adjustors 230 can be composed of the same or different materials. For instance, the stress adjustor 230 can comprise a portion of the silicon substrate 210 or an insulating material, such as silicon dioxide.

[0030] The stress adjustor 230 is positioned to modify a mobility of a majority carrier within the channel region 225. Preferably, the stress adjustor 230 is positioned within the isolation structure 215, such that a portion 235 of the isolation structure 215 is between the active area 205 and the stress adjustor 230. In some cases, it is preferable for the stress adjustor 230 to be a long continuous block. Such a shape is effective at altering the compressive stress imparted from the isolation structure 215 to the entire channel region 225. This is in contrast to active dummies, which are generally small pillar-shaped to minimize parasitic capacitances.

[0031] In some embodiments, it is advantageous for a long dimension 240 of the stress adjustor 230 to be perpendicular to an intended direction of current flow 245. In such instances, it is desirable for the long dimension 240 of the stress adjustor 230 to have a length 250 substantially equal to the gate width 228, as this facilitates modification of the mobility of a majority carrier in the entire channel region 225. However, embodiments where there

are a plurality of stress adjustors 230 between the active area 205 and isolation structure 215 are also within the scope of the present invention.

[0032] In other embodiments, it is advantageous to place the stress adjustor 230 such that its long dimension 255 is parallel to the intended direction of current flow 245. Such placements advantageously modify the extent of compressive stress imparted from regions of the isolation structure 215 that are 225 perpendicular to the intended direction of current flow 245. In such embodiments, it is desirable for the long dimension 255 to have a length 260 substantially equal to the gate length 227, as this facilitates modification of the mobility of the majority carrier in the channel region 225.

[0033] The thickness 265 and depth 270 of the stress adjustor 230 can also be adjusted to further modify the mobility of the majority carrier in the channel region 225. For example, in some embodiments, the thickness 265 of the stress adjustor 230 is least about 50 to about 500 nanometers. Such thicknesses are favored because they facilitate the ability of the stress adjustor 230 to modify the compressive stress from the isolation structure 215 to the channel region 225. In other embodiments, the depth 270 of stress adjustor 230 is substantially equal to the depth 219 of the isolation structure 215. In other cases, the depth 270 of stress adjustor 230 is substantially equal to the depth 275 of the channel

region 225. Of course, the MOS transistor 200 of the present invention can include various combinations of differently oriented and dimensioned stress adjustors 230.

[0034] The shape and choice of material comprising the stress adjustor 230 depend upon the type of MOS transistor 200 desired. For instance, when the MOS transistor 200 is an NMOS transistor, it is desirable for the stress adjustor 230 to reduce a compressive stress to the channel 225. Such embodiments are particularly advantageous when an active overlap distance 280 is small (*e.g.*, less than about 500 nanometers). Alternatively, when the MOS transistor 200 is a PMOS transistor, it is desirable for the stress adjustor 230 to increase a compressive stress to the channel 225.

[0035] It is preferable for the stress adjuster 230 to be near the active area 205 because this facilitates the stress adjustor's ability to change the mobility of the majority carriers in the channel region 225. This is contrast to active dummies which are generally distant (*e.g.*, at least about 1000 to about 2000 nanometers away) from active areas because the purpose of active dummies is to reduce dishing associated with CMP over the entire substrate surface. As noted above, the portion 235 of the isolation structure 215 is between the active area 205 and the stress adjustor 230. In some cases, a distance 282 between a perimeter 284 of the stress adjustor 230 and a perimeter 286 of the active area 205 is at least about 50 nanometers, and in some

advantageous embodiments, between about 50 and about 300 nanometers.

[0036] As indicated above, the placement of one or more stress adjustors 230 adjacent the active area 205 can substantially altered the compressive stress imparted by the isolation structure 215 to the channel region 225. Consider, for instance a MOS transistor 200 having an active overlap of about 180 nanometers and bounded by a STI 215 whose width 217 is about 600 nanometers. A silicon stress adjustor 230 is placed within the center of the STI 215. The stress adjustor 230 has a width 265 of about 200 nanometers and a depth 270 that is substantially the same as the depth 219 of the STI 215. Such a transistor has about 27 percent less compressive stress in a direction parallel with the direction of current flow 245, compared to a similar-dimensioned transistor having no stress adjustor.

[0037] Of course, the MOS transistor 200 can include other conventional device components, such as a well 290, source/drain structure 292, source/drain extension 294 and sidewalls 296, to provide an operative transistor. The presence of one or more stress adjustor 230 can substantially improve the performance characteristics of the operative MOS transistor 200, as compared to an analogous transistor with no stress adjustor. For example, certain NMOS embodiments of the MOS transistor 200, having an active overlap distance 280 of less than about 500 nanometers, have

an about 20 percent lower ratio of I_{off}/I_{on}, as compared to a similarly dimensioned transistor having no stress adjustor.

[0038] FIGURE 3 illustrates by flow diagram, another embodiment of the present invention, a process 300 for constructing an integrated circuit (IC). One skilled in the art would appreciate that the process 300 can be advantageously applied to the layout design of a variety of transistors or other semiconductor device. As noted above, this embodiment is particularly advantageous when applied to control the performance characteristics of transistors for technology nodes where the active overlap of transistors is small.

[0039] The process 300 commences, at step 305 with the generation of a mask layout for the IC. The mask layout could include planned dimensions for active areas, gates and isolation structures, for example. An active overlap distance between a planned perimeter of a selected gate and a planned perimeter of a selected active area of the IC is calculated in step 310. Those skilled in the art would be familiar with conventional mathematical algorithms can be used to facilitate this calculation.

[0040] In step 315, a compressive stress along a direction of an intended current flow through a planned channel of the IC is determined based on the active overlap distance. One skilled in the art would appreciate the variety of techniques that could be used to determine compressive stress. Such techniques could

include, for example, determining the compressive stress based on conventional theoretical understanding of the factors causing stress. Alternatively, the compressive stress can be determined with the aid of an empirical data base, where stresses have been measured in structures analogous to the instance transistor.

[0041] In step 320, it is ascertained whether or not the compressive stress determined in step 315 is greater than a critical stress parameter by comparing these values. If the compressive stress is greater, then a stress adjustor area is introduced, in step 325, adjacent the active area to modify a mobility of a majority carrier through the planned channel. If the compressive stress is not greater, then the mask design is complete, and the mask design is used to produce the IC, as indicated in step 330.

[0042] The value of the critical stress parameter can be established by theoretical or empirical techniques, similar to that described for determining the compressive stress. It is preferable to load, in step 335, different critical stress parameters for different transistor types, for use in the comparison made in step 320. As an example, the critical stress parameter preferably is adjusted to different values, depending on whether mask layout data set under consideration defines an NMOS or a PMOS transistor.

[0043] Moreover, the critical stress parameters may be different for different zones of the integrated circuit layout. Therefore,

in certain embodiments, the comparison made in step 320 is facilitated by grouping or binning together all zones of the layout that have the same critical stress parameter. Of course other parameters, such as I_{on} or transconductance (G_m) could be used as the basis for binning of zones. In other embodiments, it is advantageous to perform the process 300 on pairs of transistors that are coupled to each other in the IC. As an example, consider NMOS and PMOS transistors that are coupled together form a CMOS device. The critical stress parameter for the PMOS transistor and a critical stress parameter for the NMOS transistor could be adjusted such that the PMOS and NMOS transistors have substantially the same on-current.

[0044] In some cases, it is desirable to express the compressive stress and critical stress parameter in units other than pressure. In some instances, the critical stress parameter corresponds to a predefined critical active overlap distance and the compressive stress corresponds the active overlap distance. For example, if the active overlap for the transistor of interest is less than a critical active overlap, then that transistor is a candidate for placing a stress adjustor. One skilled in the art would appreciate that other proxies for compressive stress could be used.

[0045] In cases where the placement of a stress adjustor is indicated, it can be desirable to modify the stress adjustor, in

step 340. Any number of features of the stress adjustor can be modified to affect a change in the extent of compressive stress imparted from the isolation structure to the channel region of the transistor. For instance, increasing or decreasing one or more of the physical dimensions of the stress adjustor could modify the compressive stress. Moving the stress adjustor closer or farther from the active area could also alter the compressive stress. Changing the orientation of the stress adjustor so as to have a long dimension that is perpendicular, parallel or at intermediate angles, with respect to the intended direction of current flow through the channel, could also affect the compressive stress.

[0046] It some embodiments, it is advantageous for the modifications to the stress adjustor in step 340 in an iterative fashion. For instance, one or more of the above-described features of the stress adjustor could be modified by a predefined increment, and the determination of the compressive stress repeated in step 315, until the compressive stress is no longer greater than the critical stress parameter. During the course of this iterative process, additional stress adjustors could also be added, in step 325, if necessary.

[0047] Of course, the above-described process 300 could be applied to each active area in the IC layout. The process 300 could also include storing the mask layout data sets in temporary pattern generation (PG) files, checking for rule violations, and

repeating the process until no rules are violated. In some embodiments, these temporary PG files are then stored in a master PG file, that in turn, is used to produce mask works using convention producers. The mask works are then used for the pattern transfer of different layers comprising the IC layout to a semiconductor substrate to produce the IC. Of course, the above-described process 300 could be equally applied to the production of ICs using direct-write technologies.

[0048] Although the present invention has been described in detail, one of ordinary skill in the art should understand that they can make various changes, substitutions and alterations herein without departing from the scope of the invention.